Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

# AN ANALYTICAL SCALE FOR JUDGING HANDWRITING[1]

FRANK N. FREEMAN
University of Chicago

A number of experiences with the two handwriting scales which have been published led to the effort which is here described to construct a scale upon a somewhat different principle. The outstanding difficulty which was found with both the Thorndike and the Ayers scales is the lack of uniformity in the results which are obtained by their use. The question then arose whether some means might not be found by which this lack of uniformity might be overcome, and this led to the question as to the source of the variability in the results obtained in using the scales.

The critical statements which are made in this paper with reference to the Ayers and the Thorndike scales are not intended as a disparagement of their value for certain purposes; they are made rather for the purpose of pointing out the limitations in the use of these or similar scales. These scales are useful when it is desired to make merely a rough general survey of handwriting excellence. There is danger, however, that they may be used with an exaggerated confidence in the accuracy of results which are obtained by them. The considerations which are here given are for the purpose of preventing such uncritical confidence. It is the purpose of the paper, further, to indicate the directions in which a more reliable means of measurement is to be found.

The testimony of those who have used the handwriting scales leads to the conclusion that one important source of variability is the ambiguity as to the characteristics in writing which are to be used as the basis for judgment. Professor Thorndike, the author of one of the two widely used scales, makes the point that the use

[1] The scale itself is published in a volume entitled *The Teaching of Handwriting*, by the writer, published by Houghton Mifflin Co. This paper was read before the Section on Education of the American Association for the Advancement of Science at Philadelphia, December, 1914.

of a scale will give objectivity to the measurement of writing which cannot be had by judgment without the use of some such standard for reference. When a person attempts to judge a specimen of writing, however, by the use of the scale, the first question that arises is, What is to be taken as the basis of the judgment; what characteristic in the writing is to be looked for in order that it may be determined what part of the scale it corresponds with, and what rank is to be given to it? The answer which is given in the case of the Thorndike scale is that the judgment is to be based upon three characteristics; namely, legibility, beauty, and character.

The first difficulty with judgment upon the basis of these three characteristics is that of determining what relative weight should be given to each of the three. The judgment is a single one to be made upon the combined value of these three traits. If, now, one judge uses legibility as the chief trait in his rating, and another beauty, and another character, it will be seen that, even though each judgment may be perfectly reliable on the basis of the weight it gives to the different characteristics, yet the rating given to a particular specimen may differ widely.

Another difficulty is introduced, however, by the character of the three traits which are to be used as a basis of judgment. They themselves, even taken singly, are difficult to use in making a judgment, because each of them represents a complex characteristic which may be based on a variety of elementary characteristics in the writing itself. In brief, beauty and character and even legibility are not accurately definable characteristics so that we can be sure that different persons mean exactly the same thing by them.

A similar difficulty confronts the judge who attempts to use the Ayers scale. In this scale a distinction is not made between the different characteristics which are to be used as a basis for judgment. The term "general quality" is the only one which is given to guide the person who is to use it. Confusion arises, further, by the fact that the basis upon which the specimens for the scale were selected is not the same as that which is used by the person who attempts to grade specimens by means of it. The specimens for the scale itself were chosen on the basis of the rapidity with which ten individuals could read the specimens or other

samples of writing by the same individuals who produced the specimens in the scale. The assumption then is that there is a correlation between the rank in specimen as determined by the rate-of-reading method and the rank based upon a judgment of quality.

Merely by the examination of the specimens on the scale one is led to a suspicion that this correlation is not complete, and it is a common complaint among those who use the scale that in some cases higher specimens do not appear to be of as good quality as some of those lower down on the scale. This naturally introduces confusion in the mind of the person who is attempting to use the scale. A further confirmation of this suspicion is obtained when we conduct an experiment to determine the degree to which the correlation exists. When a number of specimens of writing are graded by the Ayers scale and then by the speed-of-reading method by the same individual, it is found that little correlation exists between the series as arranged on the basis of the two types of measurement; in fact, the coefficient is so low as to be scarcely enough to be taken as evidence of a reliable correlation. It is evident, then, that legibility is not the basis for judgment with the Ayers scale and we are left at sea as to just what the basis is.

Both for the measurement of writing and for the guide of the teacher it is very desirable that we should obtain a more exact definition of the characteristics which make up excellence in writing. If this is not done, we cannot say that objectivity is reached. The analogy with linear measurement will not hold, because in the case of linear measurement, although there may be errors in the application of a scale, yet the notion of the characteristic which is to be measured is perfectly definite. There is no ambiguity as to what is meant by *length* or by *surface area* or by *cubic contents* in the same manner as there is an ambiguity as to what is meant by *quality* or *excellence* in writing. When this ambiguity exists we are confronted with a double difficulty in the measurement of such a product as writing. We have not only the difficulty in discerning the degree of the trait in question, but also in determining just what trait it is upon which the judgment is to be based. This latter difficulty, that is, the difficulty due to ambiguity, should be

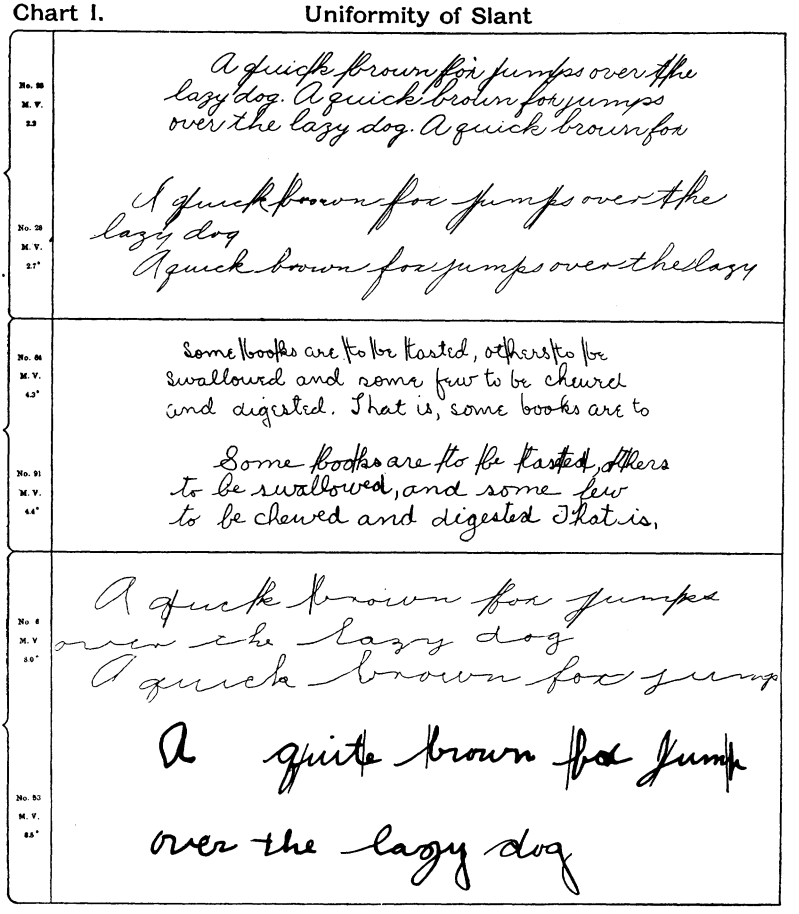
overcome so far as possible at the start. It is for the purpose of reducing, at least, this difficulty that the scale which is to be described was constructed.

When the attempt was made to determine the characteristics which should be used as the basis for judging the excellence of writing, the first one selected was uniformity. If writing is not uniform, it evidently suffers by such fact. That is, the slant of the writing, for example, should be the same throughout, and the height of those letters which are supposed to be of the same height should be equal. These illustrations suggest the two types of uniformity which can readily be determined; that is, uniformity of slant and uniformity in the alignment of the letters. These, accordingly, constitute the first two traits. The accompanying chart is a reduced copy of the first of the five charts which compose the scale. It may be referred to as an illustration of the general characteristics of the charts as a whole.

Uniformity of writing depends largely upon the fluency and the regularity of the movement by which it is produced. Another characteristic which depends also upon the ease and fluency of the movement is the quality of the line or the stroke. The line may be regular, smooth, and even, or it may be irregular, jerky, wavy in character. The one, then, constitutes excellence and the other a defect in writing. We may take quality of line, then, to be a third characteristic and one which depends largely, as in the case of uniformity, upon the character of the movement.

The movement may be smooth and regular, however, and yet the writing may be difficult to read because the letters are not well formed. We may ascribe this characteristic largely to the accuracy with which the form of the letters is perceived and the care which is taken to maintain the proper form. Letter formation has been commonly thought of as the chief form of excellence of writing. It certainly is an important one, particularly when we regard, not the deviations from an arbitrary standard of form, which are of no importance, but those deviations which make the writing difficult to read, as, for example, when one letter is made of similar form to another. The failure to round the top of the *m* or the *n* makes these letters resemble *u* or possibly *w*.

Finally, the appearance and the legibility of writing is determined very largely by the spacing between the letters or between the words or between the lines. Writing is sometimes met with which




would be very legible if it were not so closely crowded together. The writer has confirmed this by tracing the letters of such writing, but spacing them farther apart. The writing also may present a scrawled appearance because the letters of the words are spaced too widely.

This, then, gives us five characteristics which are perfectly definable in their nature and in some cases at least, as we shall see, are measurable in their degree. One problem which arises is as to the relative weight which should be given to excellence in these five characteristics determining the total grade of the writing. For the present we shall not attempt to determine accurately what this relative weight should be. We can only give them an arbitrary weight, therefore, and wait for further investigations to determine which of these is the most important in determining the total quality of the writing. We may assume, however, as a starting-point that letter formation is somewhat more important than any one of the other characteristics. In order to express this increased importance this characteristic has been given double weight in the scale.

In the first attempt to construct a scale on the basis of these five characteristics it was assumed that when the characteristics were clearly defined, series of specimens could be chosen merely upon the basis of the judgment of a group of individuals. Accordingly a preliminary chart was first constructed for the purpose of illustrating different degrees in these traits, and this was used as a guide by a class in experimental education composed of advanced students, most of whom were experienced in supervision and teaching, upon which to rank a large number of specimens into ten degrees of excellence in each trait. After the specimens had been so rated, however, it appears to the writer that the order was not always the correct one. He therefore set about to find so far as possible objective measurements upon which to base the rating and with which to compare the rating made on the basis of the judgment of the graders.

In the case of uniformity of slant and of alignment objective measurement could be directly applied. It was only necessary to measure the angle of a series of letters and to find the mean variation among these angles, or to measure the vertical position of the tops and bottoms of the letters and to measure the variability among these positions. When this was done it was found that the order based on the variability as measured did not correspond to the order on the basis of the judgment made by the graders. The

order based on objective measurement was therefore used as a basis for the selection of specimens for the scale.

In the case of quality of line no means was found upon which to base such objective measurement, but the characteristic in question was made more prominent by photographic enlargement. When this was done it was relatively easy to determine differences in the irregularity of the line of the writing. On this basis specimens for this chart were selected.

In the case of letter formation difficulty again was found, but some approach to objectivity was obtained through the assistance of Mr. R. R. Simpkins, of the State Normal School at Macomb, Illinois. Mr. Simpkins had developed a system of determining excellence in letter formation by the method of counting the errors in form. He was therefore requested to note the errors in the specimens of the scale which had been selected. In some cases the results of this measurement differed from the result of the judgment of the graders, and a compromise between the two methods of determination was used in selecting the specimens for the scale.

Some degree of objectivity in the determination of the best spacing between letters or between words was obtained in the following manner. Three widely differing styles of writing were chosen and the judgment of about fifteen people was obtained as to the most pleasing distance between the words in these three styles of writing. This was done by selecting pairs of words and placing one upon a transparent piece of paper so that it could be moved back and forth until the most agreeable distance between the words was found. This distance was then recorded by reference to a scale below the word.

It was found that this distance varied considerably with the different styles of writing, according as they differed in size or in slant. The slant, however, seemed to be the most important characteristic in determining the spacing which should be used. The length of the word seemed to have little effect in the matter. On the basis of this determination a standard spacing between words for these three styles of writing was selected. The spacing to be regarded as the standard between letters was determined in

a slightly different way. In this case various specimens were constructed in which the spacing between letters was different, and the judges were asked to select that specimen in which the letters were the most agreeably spaced.

The chart was then constructed in the following manner. Specimens were selected which should conform to the standards of spacing between letters and between words. Then the lower grades were constructed by varying this spacing in a variety of ways. The middle grade was determined by varying one of the two characteristics, that is, spacing between letters or between words, and the lowest grade by varying two of the characteristics.

Each chart contains specimens of writing which represent three grades of excellence in the characteristics in question. This small number was chosen in order that the degrees of difference might be great enough to be perfectly evident, or at least to be as evident as possible. An experiment was first made with a larger number of differences and this was not found so satisfactory. The lowest grade is given a rank of one, the middle grade the rank of three, and the highest grade the rank of five. The exception to this is in the case of letter formation, in which double weight is given to each rank. The intermediate grades between these may be used, that is, the grades of two and four, or, in the case of letter formation, of four and six. The total rank of the paper is to be determined by adding the rank in each of the five characteristics.

As has already been said, the difficulty of grading such a product as handwriting is not completely met when we have defined and represented the traits to be measured. There is still the difficulty which a judge finds in selecting the trait to be used as the basis of judgment from the complex of those characteristics which make up the writing. In order to facilitate this selection, or abstraction, various devices were used to emphasize the trait which is represented in each of the charts. In the case of uniformity of slant a line is drawn parallel to the tall letters to catch the eye and to lead to an easy determination of variability. Furthermore, a transparent paper upon which are drawn lines of a certain slant is to be placed over the specimen which is to be judged in order that the variability may be seen by comparison with these lines, all of which have the

same slant. The variability in alignment is brought out in the same way by a line above or below the letter. An effort was made to emphasize the variation in line or irregularities in line in the same way that they were emphasized for the purpose of selecting specimens for the scale, but it was found that in the reproduction of the specimens of the photographic enlargement the irregularities of the line were in a large measure smoothed out, so that this aid is of little avail. To facilitate the judgment in letter formation the errors in the first part of each specimen on the chart were indicated by small arrows. The judgment of spacing is facilitated by the artificial construction of specimens each of which illustrates a particular type of error in this respect.

It has already been remarked that all of the difficulties in using such a scale are not met when we have defined the characteristics. To place a scale before a person in which the characteristics are represented does not guarantee that he will see them. It is necessary, therefore, that some attention be paid to the training of a judge in order that he may learn to abstract from the complex written product the characteristic which is to be judged. I am convinced on the basis of my experience with a large number of persons in using scales that we cannot expect to find any device which can be used in a simple or mechanical manner or without care and some degree of preparation for the task. We do not expect people to be able to distinguish accurately between live stock or between ears of corn or any other complex object without having some training for the task. The assumption that anybody can judge handwriting or any of the other products of the school without the development of some degree of expertness will serve to amass a large number of facts which are in themselves highly unreliable and therefore to discredit the whole movement for scientific measurement of educational products.

The teacher also needs some degree of training or expertness in distinguishing the various characteristics of writing and in grading the writing of her pupils so as to determine their rank or to determine the direction in which their further efforts should be applied. In fact, the primary application of such a scale as this may be in the schoolroom where the teacher needs to distinguish, not

only differences in quality of writing, but also the particular respect in which the differences exist and the particular feature of each pupil's writing which needs to be improved.

The scale as it stands appears to be cumbersome, but after a person has acquired some skill in the application of the judgment on the various traits one may grow independent of the need of actually consulting the specimens. That is, the specimens and their rating may be kept in mind and a judgment may be made on the five characteristics with considerable rapidity.

The purpose of presenting a description of this scale for judging handwriting is not merely to make known this particular example of measuring scales, but to use it as a type of the method which can be most profitably pursued in the construction of scales in general. This method involves, first, analyzing the product into fairly simple elementary characteristics; secondly, defining and measuring these so far as possible in objective terms; and, thirdly, constructing the scale in such a way that different degrees of the elementary characteristics may be made evident to the person who is to use it.